Center on
HEALTH CARE
EFFECTIVENESS

RESEARCH BRIEF

Matching Study Designs to Research Questions in Disability-Related Comparative Effectiveness Research

July 2011

Jeffrey Ballou, Eugene Rich, and Matthew Kehn

MATHEMATICA
Policy Research

Matching Study Designs to Research Questions in Disability-Related Comparative Effectiveness Research

by Jeffrey Ballou, Eugene Rich, and Matthew Kehn

ISSUE AT A GLANCE

The impact of disability is wide ranging, influencing nearly every aspect of affected individuals' lives, including how they seek and receive health care, how they process information, and what level of independence and community participation they enjoy. Evidence-based care has potential to improve the lives of the estimated 33 million American adults with disabilities, as affirmed by the Institute of Medicine (IOM): "Disability is not destiny for either the individuals or the communities in which they live. Rather, disability is shaped by personal and collective choice. Positive choices … can mitigate the effects of [disabling conditions] and help create more supportive physical and social environments that promote a future of increased independence and integration" (Au et al. 2011; Erickson et al. 2010; Field and Jette 2007).

Comparative effectiveness research (CER) can help patients with disabilities and their providers determine which choices are beneficial and will most reduce limitations and barriers and improve quality of life. The IOM CER Priorities Committee and the Federal Coordinating Council for CER (FCCCER) both place a high priority on CER that addresses treatments and care delivery systems for people with disabilities (Ratner et al. 2009; FCCCER 2009). Well designed CER will be critical as the number of working-age adults affected by disability grows in the future, along with disability-related health expenditures, which were nearly $398 billion in 2006 (Field and Jette 2007; Anderson et al. 2011). This brief presents methodological and design issues for researchers to consider.

Challenges of Disability-Related CER

Despite the promise of high-quality disability-related CER to improve lives, the characteristics and health service needs of individuals with disabilities present challenges to the design and implementation of strong studies. The extensive heterogeneity in levels of physical and cognitive functioning, mix of limitations, and presence of different comorbidities among those with disabilities complicates the problem of recruiting adequate sample sizes and generalizing findings. Moreover, the range of interventions for people with disabilities is extremely broad—ranging from traditional medical and surgical treatments to supported employment, home modifications, and peer mentoring programs. It is not obvious that all of these types of interventions will necessarily require the most rigorous randomized controlled trials (RCTs); indeed, many interventions related to care coordination and the provision of home and community-based services (HCBS) are either infeasible or unethical to test in an RCT (Johnston et al. 2009).

One consequence of the challenges of conducting strong CER studies is that individuals with disabilities, providers, and policymakers often do not have access to the evidence that they need to decide whether to participate in a program, choose an intervention, or

implement a policy. Absent appropriate evidence, decision makers often turn to anecdotal information or opinions from the field for guidance. Although such information can be valuable, the availability of high-quality CER based on well-established research principles would do much to bolster the confidence that decision makers place in their decisions. If only RCTs—which can be time-consuming and expensive—can provide sufficiently strong evidence for most disability-related CER questions, this dearth of evidence is likely to persist. If for certain questions it is reasonable to enlarge the set of study designs that can provide useful information, however, the possibilities for developing and evaluating evidence to inform decision making expand considerably.

More generally, the Methodology Committee of the Patient-Centered Outcomes Research Institute (PCORI), established under the Patient Protection and Affordable Care Act, has been charged with developing a "translation table" to match CER questions with appropriate research designs. This high-profile initiative represents an acknowledgement of the potential value of traditional RCTs but also other, alternative study designs in contributing to the availability of more and better CER that provides timely and informative evidence to the decision makers who require it. As a broader variety of study designs come to be viewed as acceptable for addressing certain disability-relevant CER questions, decision makers will have to use judgment in assessing which studies and study designs provide the required level of evidence and therefore may be included in the body of evidence to be evaluated. The purpose of this report is to provide decision makers with guidance on how to exercise that judgment.[1]

What Evidence Is Needed?

Although extensive criteria for evaluating the quality of research evidence have been developed, they provide only limited guidance to decision makers on which study designs are appropriate for answering different types of CER questions for adults with disabilities. Therefore, controversy remains regarding how to apply these criteria in the context of the real-world questions to which decision makers seek answers. One former Medicaid director noted that although the answers to some questions require large RCTs, the sort of descriptive study done by *Consumer Reports* also can offer valuable information on, for example, the battery life of different powered wheelchairs. The breadth of concerns and relevant interventions, combined with the wide heterogeneity among this population, suggests that different methodological approaches will be appropriate for different interventions targeting different populations. The challenge becomes one of determining which study designs are informative in a given context.

Decision makers reviewing evidence often seek answers to a particular question regarding the effectiveness of a specific intervention or service for individuals with a specific type of disability (see box on p. 3). To use existing evidence most effectively, they must first determine the required levels of internal validity (confidence that the intervention indeed caused the observed benefits) and external validity (generalizability of findings beyond the study's sample) (Table 1). It then becomes possible to ask which study designs are acceptable for providing the needed information, recognizing that in many cases the ideal study will not be

[1]Although the principles developed here are intended to help decision makers evaluate research, it is our hope that researchers—anticipating how their research will be assessed by those who use it—will also find these guidelines to be helpful as they seek to conduct studies that will most effectively drive evidence-based change and improve the lives of people with disabilities.

available and trade-offs regarding how to use available evidence will be required. Finally, the set of identified studies must be evaluated using established and accepted standards of evidence, such as those promulgated by the Cochrane Collaboration and others. Although this approach to using evidence as the basis for decision making is applicable to CER generally, each step has specific implications for disability-related research.

EXAMPLES OF REAL-WORLD CER QUESTIONS

- What is the comparative effectiveness relative to usual care of a new medication to ameliorate the effects of Alzheimer's disease?
- What is the relative effectiveness of powered versus manual wheelchairs for improving the independence of working-age adults with physical disabilities and no other impairments?
- Does implementing an advanced medical home model coordinating medical and HCBS services for adults with disabilities result in lower adjusted rates of adverse outcomes, such as nursing home admissions and emergency room visits, compared with traditional primary medical care and HCBS programs?
- Are individuals with disabilities who participate in a peer mentoring program less likely to develop secondary conditions than individuals who do not participate?

1. When Is High Internal Validity Required?

High internal validity, requiring a rigorous study design such as an RCT, generally will be required whenever an intervention is costly to implement, has the potential to cause significant and/or irreversible harm, or both (Zietman 2010). Although many interventions relevant to disability services—such as the provision of HCBS or care coordination strategies—likely will not be associated with a high potential for causing direct harm that an untested drug might carry, the scope and potential cost of such large-scale disability-related interventions may demand evidence with high internal validity.

Costly interventions. Costly interventions may require high internal validity because the opportunity costs to the health care system associated with being wrong rise with the intervention's cost.[2] For example, there might be substantial incremental costs related to the use of new pharmaceuticals, implantable devices, or assistive technologies that must be covered by raising premiums, limiting other benefits, or increasing out-of-pocket payments. Interventions involving major reorganizations either within or across medical practices or other organizations serving people with disabilities (for example, the conversion of practices to patient-centered medical homes—or the bundling of HCBS care coordination with in-home geriatric assessments—can have substantial additional costs as well. These might require extensive retraining of medical professionals and/or modifications of infrastructure and care processes, changes demanding substantial professional time and other organizational resources that could have been applied elsewhere.

[2]The typically high cost of RCTs is also easier to justify for costly interventions. Policymakers and funding organizations might be more likely to support expensive research projects with high internal validity in order to avoid spending an even larger amount on an inadequately tested and ultimately ineffective costly innovation.

Possibility of significant harm. Whenever an intervention otherwise intended to improve health nonetheless has the potential to cause serious and/or irreversible harm—or when the uncertainty surrounding assessments of potential harm from an intervention is sufficiently high—a high level of internal validity is typically required to avoid the risk of acting on incorrect evidence. For example, given the range of current safe and effective therapies for depression, CER on a new drug likely would require the high level of internal validity offered by a well-designed RCT.[3] Harm from interventions can be direct or indirect. For example, a new home monitoring technology for adults with intellectual disability might reduce the need for active monitoring by caregivers; but the risk of harm from the sudden unpredictable failure of the intervention might suggest the need for studies with high internal validity to assess safety and effectiveness.

2. When Might Evidence with Lower Internal Validity Be Informative?

Decision makers might not require evidence of the highest possible level of internal validity for interventions in which both the possibility of significant harm and the cost of treatment are low. Similarly, an intervention that slightly modifies an existing approach already supported by strong evidence of effectiveness could be assessed using designs that provide lower internal validity, such as observational studies. Indeed, when the modifications are sufficiently specific and inexpensive, the requisite evidence could be the case-by-case assessment of individual consumers. For example, a decision about the best wheelchair modifications for an individual with a particular mix of limitations will be driven by the user's specific capabilities and can likely be addressed through simple observational studies.

Decision makers might choose to employ an intervention even in the absence of evidence with high internal validity when the current standard of care for the condition targeted by the intervention is virtually certain to result in imminent, serious, and irreversible harm. Thus, newly available treatments for conditions that typically result in severe disability or death in a short time might be elected on the basis of relatively weaker evidence whenever the results of an RCT are not yet available.

3. What Level of External Validity Is Required?

One challenge of conducting good disability-related CER is designing studies that will generate findings that are applicable (that is, externally valid) beyond the study's sample. Interventions that can be applied to a diverse population in a variety of different environments or practice settings will require high external validity. Many observational studies have high external validity: most observe treatments administered in real-world environments and many are designed specifically to address either a broad set of subgroups or subgroups not previously covered in trials. Although RCTs may also have high external validity, the extent of applicability can be limited by selection in the participation decision[4] and the high cost of procuring the sample size needed to observe effects in multiple subgroups of patients and settings.

---

[3]Assessing the potential for harm a priori can be extremely difficult, depending on the nature of the intervention, and actually might be aided by observational evidence that points to the need for a trial. Moreover, significant harm to the population administered the intervention could occur because of a high probability of negative effects, a high severity of negative effects, or both. Even large and lengthy RCTs might fail to properly identify the extent of harm caused by a newly developed drug if the probability of negative effects is sufficiently low (that is, when outcomes involving harm are rare).

[4]See, for example, Tamer and Kline (2011) and Heckman (1992).

Interventions applied to a diverse range of people with disabilities. Individuals with disabilities differ in many ways, including the ability to perform activities of daily living, the ability to fully understand and participate in a course of treatment, the presence of multiple chronic conditions, and the residential and community settings in which they live. Given this diversity, decision makers considering an intervention aimed at a broad set of individuals with disabilities will require high external validity addressing effectiveness in the diverse subgroup populations of interest. For example, a study of the effectiveness of community group homes for individuals with mental illness might be of limited value to a decision maker whose population of interest is significantly more or less severely ill, on average, than the study's population.

Decision makers might also demand high external validity when considering interventions applicable to the general population. If the intervention includes elements of care delivery that might present difficulties for individuals with certain impairments—such as traveling twice weekly to a clinic for treatment—the decision maker should ask whether the intervention was tested on a subgroup of individuals with disabilities and what the nature of those disabilities was.

Table 1. Criteria for Determining Needed Levels of Internal and External Validity

Higher Internal Validity	Higher External Validity
Intervention is costly to implement OR Intervention might cause significant harm	Intervention is for diverse populations in different settings OR Implementation will be difficult to replicate OR Intervention has to be tailored to specific settings
Lower Internal Validity	**Lower External Validity**
None of the criteria for highest internal validity are met AND/OR Imminent risk of harm if intervention is not implemented	Implementation can be replicated in different settings AND Coordination across/within organizations is *not* required AND Uncomplicated for patients to participate in treatment

Source: Mathematica Policy Research.

Complex or customized interventions. Decision makers likely will require evidence with high external validity for complex interventions that are difficult to implement faithfully in typical practice settings. Examples of such interventions include those that require extensive retraining, coordination within and across practices, or substantial participation on the part of the individual being treated. High external validity also will be required if an

intervention has to be tailored to a specific environment. For example, supported employment interventions can require extensive interaction among the case manager, the individual with a disability, the work incentives counselor, and the potential employer. These interventions might be implemented differently for each individual seeking support, suggesting that decision makers should exercise caution in applying findings from a study that analyzes a narrowly defined population or a single environment.

Studies of interventions that are straightforward to administer in a variety of settings and require little coordination by providers and caregivers within and across organizations—and relatively low levels of patient participation—might generalize well beyond the study's sample. We expect that such circumstances are relatively rare in the world of disability-related CER, however.

Which Study Designs Are Informative and What Trade-Offs Have to Be Made?

When the decision maker has clarified the levels of internal and external validity that are needed, it is possible to determine which studies might be most informative to guide the choices at hand. Often, the ideal study will not exist, requiring the decision maker to exercise judgment in deciding whether and how to use the available evidence.

1. Which Study Designs Are Informative?

As reflected in standards published by the Cochrane Collaboration and others, a need for the highest internal validity will generally require a strong RCT or pragmatic clinical trial (PCT),[5] because proper randomization guarantees the validity of the estimator's statistical properties, including the comparability, on average, of treatment and comparison groups at baseline and during follow-up.[6] However, rigorous observational studies might be sufficient if findings with a moderate level of internal validity will be useful to decision makers, especially if RCTs cannot provide the required degree of external validity.

Because testing interventions in real-world settings is especially important for disability-related research, PCTs are ideal whenever high external validity is required. The number of available and relevant PCTs likely will be limited, however, because it is expensive to conduct strong PCTs with sizable numbers of different subgroups. Service providers and policymakers also may consider carefully executed observational studies that examine relevant existing data and have strong external validity. If the data are sufficiently detailed, these studies can control for the presence of comorbidities and differing degrees of impairment, allowing application of findings to a broader set of individuals.

2. Which Study Designs Are Available?

There is often little or no evidence available for making a treatment or coverage decision for people with disabilities. If studies have been done, decision makers have to determine whether they provide the necessary level of internal validity and applicability to the popula-

[5]PCTs increase the external validity of traditional RCTs by recruiting participants who would be likely to receive the intervention in the real world and administering the intervention in real-world settings.

[6]Recently, the What Works Clearinghouse has indicated that strong regression discontinuity designs may also be graded as meeting evidence standards without reservations (Schochet et al. 2010).

tion of interest. When the highest possible level of internal validity is required, this amounts to determining whether there are relevant, well-executed RCTs or PCTs.

For HCBS and care coordination interventions, RCTs will often be unethical or infeasible. For example, testing the effectiveness of community group homes for individuals with developmental disabilities, relative to an alternative such as care at home, would require randomly selecting individuals for the treatment group and moving them out of their current homes. Obtaining informed consent from all participants could present an ethical obstacle to recruitment depending on the severity of impairment of the target population. If the RCT is not ethical, its feasibility becomes a moot point.

If an RCT or PCT is ethical, it still might not be feasible for a variety of reasons. As suggested earlier, the cost of conducting a proper PCT with sufficiently high external validity might be prohibitive due to the large number of subgroups that a credible disability study might require. Moreover, there might not be sufficient time for follow-up or the decision maker might require an answer before the study can be completed.

Feasibility can also be a concern with observational studies. One of the most common difficulties with retrospective observational designs is that they do not generate sufficiently relevant or detailed data to be useful in answering the question of interest. For example, detailed data on the use of a newly developed assistive technology might not be immediately available to researchers, precluding the possibility of an observational study. In some cases, moreover, an observational approach (for example, assessing the effectiveness of different supported employment models) might require data from numerous sources, some of which (such as the employer's) could be proprietary.

3. Managing Trade-Offs Effectively

Because studies with the ideal design often will not be available, decision makers will confront trade-offs when making decisions based on less than ideal evidence. The difficulty of trade-offs is especially evident whenever the decision is one that ideally requires evidence of both high internal and external validity. If findings from credible and broadly applicable PCTs are available, little compromise is required, but if the evidence consists of both traditional RCTs (with high internal but lower external validity) and observational studies (perhaps with high external but lower internal validity), the need for trade-offs becomes explicit whenever different classes of study design report markedly different findings. In such a circumstance, the decision maker's judgment will be important in determining whether to give higher weight to studies with the highest internal validity or those with high applicability to the decision maker's population.

In other circumstances, the highest possible internal validity might be required, but evidence from strong RCTs is not available. This will frequently be true for CER on services and delivery strategies for people with disabilities, as randomization might be unethical or infeasible. In some cases, randomization will be feasible but the intervention might be so obviously different from the standard of care that administrators and/or participants will be aware of treatment assignment—that is, the experiment will not be properly blinded. In these circumstances, the decision maker must exercise judgment in deciding whether to consider evidence from weaker RCTs, nonrandomized trials, cohort studies, and other designs. Because none of these offer the level of internal validity provided by strong RCTs,

the decision maker must consider carefully the consequences of making a decision based on incorrect evidence.

Finally, even when an RCT is both ethical and feasible, the decision maker's time frame often will require a decision before the study can be completed, leaving two options: either introduce the intervention and risk incurring costs or adverse consequences that might outweigh the actual benefits of the intervention or delay providing an intervention that might prove beneficial. One potential solution is to make provisional judgments based on existing observational research, if available, while awaiting results of studies with the requisite internal or external validity. This is one of the circumstances in which CER policy experts have proposed to employ "coverage with evidence development," thus ensuring access to promising interventions but requiring collection of additional evidence employing more robust research designs to ensure that future use of the intervention is evidence-based (Miller and Pearson 2008; Tunis and Pearson 2006).

Some Examples

To better illuminate the connection between these principles for study selection and issues that decision makers could reasonably be expected to face, we briefly consider some real-world CER questions. Because a thorough evaluation of each of these would require a systematic review of the available evidence and, in many cases, careful estimation of rates of side effects and magnitudes of potential benefit, this discussion is for illustrative purposes only.

- *"What is the comparative effectiveness relative to usual care of a new medication to ameliorate the effects of Alzheimer's disease?"*

  Providers are concerned that the potential benefits of the drug for their patients might be outweighed by possible harms and patients' out-of-pocket costs for the drug. The policymaker is concerned about not only the potential net benefit of the drug for all covered patients but also the costs to the health plan and effects on premiums and other benefits. Thus, a study of high internal validity such as an RCT typically is warranted. If appropriate use of the new drug is not complicated for providers or caregivers, then decision makers will have less need for studies showing external validity.

- *"What is the relative effectiveness of powered versus manual wheelchairs for improving the independence of working-age adults with physical disabilities and no other impairments?"*

  Here the intervention has negligible obvious risk of doing harm to the individual; its cost implications might be complex or straightforward depending on the insurance status, economic circumstances, and employability of the population of interest. Proper use of the wheelchair will not be complicated for the subjects of this intervention. Thus, in many cases the evidence needed by decision makers can be of relatively modest internal and external validity.

- *"Does implementing an advanced medical home model coordinating medical and HCBS services for adults with disabilities result in lower adjusted rates of adverse outcomes, such as nursing home admissions and emergency room visits, compared with traditional primary medical care and HCBS programs?"*

The direct harms to individuals might again be negligible, but the potential societal costs of substantial medical home payments and expanded HCBS service provision could be large. Further, proper implementation of the intervention could be complex and multi-faceted. Accordingly, a study design conferring both higher internal and external validity might be needed to guide decision makers. Thus, a PCT or a sophisticated quasi-experimental design (depending on ethics and feasibility) would be needed.

- *"Are individuals with disabilities who participate in a peer mentoring program less likely to develop secondary conditions than individuals who do not participate?"*

  For this intervention, both the risks to subjects and the costs to provider organizations are likely to be negligible. To the extent that the intervention prevents the development of secondary conditions, the program might even lead to savings. However, this intervention could be complicated to implement because it involves diverse pairings of adults with disabilities and peer mentors. Thus, a research design of modest internal validity but high external validity might be most informative to provide evidence of effectiveness across a range of communities and consumer–peer mentor dyads.

Of note, it seems likely that answering many CER questions relevant to disability services will require heavier reliance on observational studies. Inevitably, the question will arise of which *observational* designs are most appropriate for a given research question, when the decision maker concludes that observational evidence can be informative. Although we do not address this issue here, we believe it is important and will require careful examination in the future.

Evolving Study Selection and Evaluation Principles

The issue of which study designs are informative to specific research questions is the subject of ongoing debate among researchers and policymakers interested in CER; several separate efforts are underway to provide both researchers and decision makers with guidance on selecting study designs either when beginning a study or when seeking the strongest and most relevant evidence to review. These include the charge to the PCORI's Methodology Committee to develop a "translation table that is designed to provide guidance and act as a reference … to determine research methods that are most likely to address each specific research question" (U.S. Public Law 111–148, Sec. 301) and projects and symposiums convened by the Agency for Healthcare Research and Quality, the National Institutes of Health, and others.[7] In the coming months and years, these initiatives will help to clarify which study designs are suitable for different subjects and how these different designs—both trials and observational—can be improved.

The acceptability of observational studies and other alternatives to RCTs for addressing certain disability-related CER questions does not relieve decision makers of the responsibility to exercise sound judgment in evaluating the relative strength of existing studies, but rather increases it by requiring them to assess *which* specific observational studies (or RCTs) are sufficiently strong to be included in the body of evidence that will inform the decision. Consequently, it is important for decision makers to pay close attention to existing

[7]See, for example, http://www.pcori.org/; http://www.effectivehealthcare.ahrq.gov/index.cfm/methodological-challenges-in-cer-conference/; and http://www.academyhealth.org/Programs/content.cfm?ItemNumber=5666&navItemNumber=2933.

standards for evaluating research.[8] Nevertheless, given the challenges of conducting strong RCTs for many questions related to services and care delivery for people with disabilities, the production and acceptance of strong evidence from alternative study designs would be a welcome development.

METHODS

To better address the need for high-quality and informative research on the effectiveness of interventions for people with disabilities, the U.S. Department of Health and Human Services asked Mathematica to develop standards for the conduct of disability-related CER and the selection of relevant outcomes measures. The guidance for selecting study designs presented in this brief was developed after a careful review of existing standards developed by prominent agencies and organizations that conduct, report, or review CER. Included in this review were producers of systematic reviews, such as the Agency for Healthcare Research and Quality, the National Institutes of Health, and the Cochrane Collaboration; clinical guideline developers such as the U.S. Preventive Services Task Force and IOM; developers of evidence assessment instruments such as the GRADE Working Group; and policymakers, such as the National Institute for Health and Clinical Excellence in the United Kingdom. Existing criteria were considered for their applicability to disability-related CER and a synthesis of findings was presented to three technical expert panels (TEPs), which consisted of CER experts, disability researchers, consumer advocates, policymakers, and program administrators. These TEPs were instrumental in affirming the accuracy of our synthesis, identifying additional criteria, highlighting the unique needs of disability-related CER, and describing how disability-related CER standards could be used in a real-world setting.

[8]Although there are well established standards for evaluating traditional RCTs and some observational designs, the increasing visibility and adoption of alternative methods, including Bayesian/adaptive, N-of-1, propensity score matching, and instrumental variable techniques, has spurred initiatives to produce new or updated standards for evaluating studies from this broader array of designs.

References

Anderson, Wayne L., Joshua M. Wiener, Eric A. Finkelstein, and Brian S. Armour. "Estimates of National Health Care Expenditures Associated with Disability." *Journal of Disability Policy Studies*, vol. 21, no. 4, 2011, pp. 230–240.

Au, Melanie, Samuel Simon, Arnold Chen, Debra Lipson, Gilbert Gimm, and Eugene Rich. "Comparative Effectiveness of Care Coordination for Adults with Disabilities," Research brief, Washington, DC: Mathematica Policy Research, 2011.

Erickson, W., C. Lee, and S. Von Schrader. "2008 Disability Status Report: The United States." Ithaca, NY: Cornell University Rehabilitation Research and Training Center on Disability Demographics and Statistics, 2010. Available at http://www.ilr.cornell.edu/edi/disabilitystatistics/. Accessed on July 30, 2010.

Federal Coordinating Council for Comparative Effectiveness Research. "Report to the President and Congress." Washington, DC: U.S. Department of Health and Human Services, June 2009. Available at http://www.hhs.gov/recovery/programs/cer/cerannualrpt.pdf. Accessed on August 9, 2010.

Field, Marilyn J., and Alan M. Jette. "The Future of Disability in America." Report by the Institute of Medicine. Washington, DC: National Academies Press, 2007. Available at http://www.nap.edu/catalog/11898.html. Accessed on September 2, 2010.

Heckman, James J. "Randomization and Social Policy Evaluation." In *Evaluating Welfare and Training Programs*, edited by Charles F. Manski and Irwin Garfinkel. Cambridge, MA: Harvard University Press, 1992.

Johnston, Mark V., Gregg C. Vanderheiden, Marianne D. Farkas, E. Sally Rogers, Jean Ann Summers, and John D. Westbrook. "The Challenge of Evidence in Disability and Rehabilitation Research and Practice." Austin, TX: SEDL, 2009. Available at http://www.ncddr.org/kt/products/tfpapers/tfse_challenge/tfse_challenge.pdf. Accessed on June 10, 2011.

Miller, Franklin G., and Steven D. Pearson. "Coverage with Evidence Development: Ethical Issues and Policy Implications." *Medical Care*, vol. 46, no. 7, 2008, pp. 746–751.

Ratner, Robert, Jill Eden, Dianne Wolman, Sheldon Greenfield, and Harold Sox, eds. "Initial National Priorities for Comparative Effectiveness Research." Washington, DC: National Academies Press, 2009. Available at http://www.iom.edu/Reports/2009/ComparativeEffectivenessResearchPriorities.aspx. Accessed on August 9, 2010.

Schochet, P., T. Cook, J. Deke, G. Imbens, J. R. Lockwood, J. Porter, and J. Smith. "Standards for Regression Discontinuity Designs." Washington, DC: What Works Clearinghouse, 2010. Available at http://ies.ed.gov/ncee/wwc/pdf/wwc_rd.pdf. Accessed on June 10, 2011.

Tamer, Elie T., and Brendan Kline. "Using Observational Versus Randomized Controlled Trial Data to Learn About Treatment Effects." Evanston, IL: Northwestern University, 2011. Available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1810114. Accessed on June 10, 2011.

Tunis, Sean R., and Steven D. Pearson. "Coverage Options for Promising Technologies: Medicare's 'Coverage with Evidence Development.'" *Health Affairs*, vol. 25, no. 5, 2006, pp. 1218–1230.

U.S. Public Law 111–148. 111th Congress, 2d. sess., March 23, 2010. *Patient Protection and Affordable Care Act*.

Zietman, Anthony. "CER in Prostate Cancer: The View of the Radiation Oncology Community." Presentation at the National Institutes of Health and Agency for Healthcare Reform and Quality Conference on Methodological Challenges in Comparative Effectiveness Research, December 2, 2010. Available at http://www.conferences.thehill-group.com/CERDecemberConference/presentations/Day%201/Zietman.Anthony.pdf. Accessed on January 16, 2011.

For more information, contact Jeffrey Ballou, senior researcher, at jpballou@mathematica-mpr.com.

This brief is based on work funded by the U.S. Department of Health and Human Services.

Center on
HEALTH CARE
EFFECTIVENESS

Improving public well-being by conducting high-quality, objective research and surveys

Princeton, NJ • Ann Arbor, MI • Cambridge, MA • Chicago, IL • Oakland, CA • Washington, DC

MATHEMATICA
Policy Research

www.mathematica-mpr.com